ITV, mid-ventilation, gating or couch tracking - A comparison of respiratory motion-management techniques based on 4D dose calculations

Ehrbar, Stefanie ; Jöhl, Alexander ; Tartas, Adrianna ; Stark, Luisa Sabrina ; Riesterer, Oliver ; Klöck, Stephan ; Guckenberger, Matthias ; Tanadini-Lang, Stephanie

Abstract: PURPOSE: Respiratory motion-management techniques (MMT) aim to ensure tumor dose coverage while sparing lung tissue. Dynamic treatment-couch tracking of the moving tumor is a promising new MMT and was compared to the internal-target-volume (ITV) concept, the mid-ventilation (MidV) principle and the gating approach in a planning study based on 4D dose calculations. METHODS: For twenty patients with lung lesions, planning target volumes (PTV) were adapted to the MMT and stereotactic body radiotherapy treatments were prepared with the 65%-isodose enclosing the PTV. For tracking, three concepts for target volume definition were considered: Including the gross tumor volume of one phase (single-phase tracking), including deformations between phases (multi-phase tracking) and additionally including tracking latencies of a couch tracking system (reliable couch tracking). The accumulated tumor and lung doses were estimated with 4D dose calculations based on 4D-CT datasets and deformable image registration. RESULTS: Single-phase tracking showed the lowest ipsilateral lung Dmean (median: 3.3Gy), followed by multi-phase tracking, gating, reliable couch tracking, MidV and ITV concepts (3.6, 3.8, 4.1, 4.3 and 4.8Gy). The 4D dose calculations showed the MidV and single-phase tracking overestimated the target mean dose (-2.3% and -1.3%), while it was slightly underestimated by the other MMT (<+1%). CONCLUSION: The ITV concept ensures tumor coverage, but exposes the lung tissue to a higher dose. The MidV, gating and tracking concepts were shown to reduce the lung dose. Neglecting non-translational changes of the tumor in the target volume definition for tracking results in a slightly reduced target coverage. The slightly inferior dose coverage for MidV should be considered when applying this technique clinically.

DOI: https://doi.org/10.1016/j.radonc.2017.05.016

couch tracking - A comparison of respiratory motion-management techniques based on 4D dose calculations. Radiotherapy and Oncology, 124(1):80-88.
DOI: https://doi.org/10.1016/j.radonc.2017.05.016

Received Date: 11 November 2016
Revised Date: 19 April 2017
Accepted Date: 20 May 2017

Title page

Proposed journal: Radiotherapy and Oncology

Title: ITV, mid-ventilation, gating or couch tracking – A comparison of respiratory motion-management techniques based on 4D dose calculations

Authors: Stefanie Ehrbar[+1,2], Alexander Jöhl[1,3], Adrianna Tartas[1,4], Luisa Sabrina Stark[1,2], Oliver Riesterer[1,2], Stephan Klöck[1,2], Matthias Guckenberger[1,2], Stephanie Tanadini-Lang[1,2]

Addresses:
[1] *Department of Radiation Oncology, University Hospital Zurich (USZ), Switzerland*
[2] *University of Zurich, Switzerland*
[3] *Product Development Group Zurich, Department of Mechanical and Process Engineering, ETH Zurich, Switzerland*
[4] *University of Warsaw, Faculty of Physics, Poland*

[+] *Corresponding author:* *Department of Radiation Oncology*
University Hospital Zurich
Rämistrasse 100
CH-8091 Zurich
+41 (0)44 255 1808
stefanie.ehrbar@usz.ch

Key words: Lung cancer, Tracking, Gating, Respiratory motion management, Stereotactic body radiation therapy, 4D dose calculation

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as
doi: 10.1016/j.radonc.2017.05.016

Abstract

Purpose: Respiratory motion-management techniques (MMT) aim to ensure tumor dose coverage while sparing lung tissue. Dynamic treatment-couch tracking of the moving tumor is a promising new MMT and was compared to the internal-target-volume (ITV) concept, the mid-ventilation (MidV) principle and the gating approach in a planning study based on 4D dose calculations.

Methods: For twenty patients with lung lesions, planning target volumes (PTV) were adapted to the MMT and stereotactic body radiotherapy treatments were prepared with the 65%-isodose enclosing the PTV. For tracking, three concepts for target volume definition were considered: Including the gross tumor volume of one phase (single-phase tracking), including deformations between phases multi-phase tracking) and additionally including tracking latencies of a couch tracking system (reliable couch tracking). The accumulated tumor and lung doses were estimated with 4D dose calculations based on 4D-CT datasets and deformable image registration.

Results: Single-phase tracking showed the lowest ipsilateral lung Dmean (median: 3.3 Gy), followed by multi-phase tracking, gating, reliable couch tracking, MidV and ITV concept (3.6, 3.8, 4.1, 4.3 and 4.8 Gy). The 4D dose calculations showed the MidV and single-phase tracking overestimated the target Dmean (-2.3% and -1.3%), while it was slightly underestimated by the other MMT (<+1%).

Conclusion: The ITV concept ensures tumor coverage, but exposes the lung tissue to a higher dose. The MidV, gating and tracking concept were shown to reduce the lung dose. Neglecting non-translational changes of the tumor in the target volume definition for tracking results in a slightly reduced target coverage. The slightly inferior dose coverage for MidV should be considered when applying this technique clinically.

Introduction

Respiratory tumor motion increases the position uncertainty in radiotherapy. Lung tumors were reported to move up to 34 mm in cranial-caudal direction [1]. To achieve local tumor control, respiratory motion has to be addressed in the treatment planning and delivery. This is commonly done by increasing the treatment volume to cover the whole extent of motion. Advanced motion-management techniques (MMT) aim to further mitigate the tumor motion, eventually resulting in a reduction of treatment volume, associated with reduced healthy tissue irradiation.

In stereotactic body radiotherapy (SBRT), the total radiation dose is delivered to the patient in only a few high-dose fractions with steep dose gradients outside the treatment volume. In contrast to conventional therapy, lung SBRT shows reduced toxicities and better quality of life [2]. Despite already low toxicities, some patient groups could further benefit from a reduction in treatment volume. Oligo-metastatic patients often get multiple courses of radiation therapy, and therefore the cumulative organ dose has to be kept as low as possible. Also for patients with large sized or centrally located lung tumors, the organ dose could be reduced with MMT in order to make SBRT with curative intent at all feasible.

The prominent motion-management techniques under free breathing are the internal-target-volume (ITV) concept, the mid-ventilation (MidV) principle, respiratory gating, and dynamic target tracking. The ITV concept [3] considers the tumor motion with a motion-encompassing safety margin. The MidV concept [4] considers the motion as random positioning error in a probabilistic safety margin calculation, leading to smaller treatment volumes than the ITV concept. Both concepts are passive motion-management techniques, which require only the a priori extent of tumor motion. Gating and tracking are active techniques, which require information on the tumor position in real-time. The gating concept [5, 6] restricts the irradiation to a predefined tumor position window, reducing the treatment volume, but increasing the overall treatment time. The tracking concept allows for volume reduction with continuous irradiation using real-time compensation of the tumor motion. Tracking can be realized by either following the motion with the treatment beam [7, 8, 9], or by counter-steering the patients, using the robotic couch, in real-time according to their internal tumor motion [10, 11, 12, 13].

For a comparison of these techniques, an estimation of the healthy tissue doses, as well as an estimation of the target doses are required. The accumulated dose to a moving tumor will differ from the dose estimation from a three-dimensional treatment plan due to interplay, gradient and density effects. The gradient effect is the blurring of the dose when the tumor moves through an inhomogeneous dose field. The temporal interplay between dynamic treatment delivery and tumor motion changes the occupation time of different tumor regions in the high dose region and causes the regions to receive higher or lower doses than expected. Density effects are caused by the respiration-

induced density variations, yielding different dose distributions in different respiratory phases. Additionally, the density of a respiratory phase differs from the density of the average CT reconstruction, which is often used for dose calculation. Therefore, this would yield a different dose to the tumor.

All of these effects can be considered in four-dimensional (4D) dose calculations to better estimate the dose to a moving tumor by a recalculation of time-resolved treatment plans on a 4D-CT image set combined with deformable dose summation. This has already been done in previous studies for ITV, MidV and tracking concepts [14, 15, 16, 17, 18, 19]. The established differences were considered to be negligibly small. However, treatment planning was performed with a homogeneous prescription to the PTV. SBRT is often prescribed inhomogeneously to the 60-80% isodose-line. This leads to higher dose gradients within the target which may lead to higher sensitivity of the accumulated dose to motion effects, which has not yet been investigated.

Planning [20, 21, 7, 17, 18] and phantom studies [22, 23] on comparison of MMT have been performed. But so far, no study has compared all four techniques in a consistent study design while considering tumor coverage and lung dose sparing as endpoints.

This planning study was designed to evaluate the benefits and drawbacks of ITV, MidV, gating and couch tracking concepts for lung SBRT. 4D dose calculations were performed to investigate not only reductions in lung doses but also target dose coverage and thus to evaluate the robustness of the treatment concepts against respiratory motion.

Material and Methods

Patients

Twenty patients with thoracic lesions were selected based on their tumor motion, having a minimal motion of 5 mm and at least half of the patients were requested to have motion larger than 10 mm. All 4D-CT scans were taken on a SOMATOM Definition AS Open (Siemens AG, Germany) CT scanner prior to treatment. Ten respiratory phases were reconstructed based on the respiratory signal of the Real-time Position Management System (Varian Medical System, Palo Alto, CA, USA). The respiratory phases were phase-binned (13 patients) or amplitude-binned (7-patients). The tumor motion was calculated as the maximal displacement of the center of volume (COV) of the tumor in all respiratory phases. Patient characteristics as diagnosis, respiratory period, GTV size and tumor motion amplitude are listed in the *Supplement*.

Contouring

The GTV was delineated by a radiation oncologist in a 4D-CT phase between inhale and exhale (initial contouring phase). The GTV contour was then propagated to the other phases with deformable image

registrations performed with the MIM Software (MIM Software Inc., Cleveland OH, USA). No extensions from GTV to clinical target volume were added. The GTVs of the ten phases were used for the definition of the radiotherapy target volumes (TV) for six treatment concepts. The TV were contoured in the following ways:

- *ITV concept*: The union of all GTVs was taken as internal target volume (ITV).
- *MidV*: The time-weighted average of the tumor positions was calculated and the GTV closest to it was defined as mid-ventilation TV [4, 24, 25].
- *Gating*: The union of the GTVs within a 30%-duty cycle around end of inhale were combined to a gating TV.
- *Single-phase Tracking*: The TV was restricted to only the GTV in the MidV-phase. This single-phase approach neglects any target deformation and rotation or residual compensation errors at the planning stage.
- *Multi-phase Tracking:* To additionally consider non-translational changes of the target volume, all ten GTVs were overlaid according to their COV and the union was taken as TV.
- *Reliable Couch Tracking*: Residual tracking errors (see *below*) of a couch tracking system and non-translational changes were considered in this approach. The ten GTVs were also overlaid according to their COV, but to include the residual tracking errors, the COV shifted apart according to the residual motion amplitude and then combined to the TV.

Based on these target volumes, the planning target volumes (PTV) were derived by adding a margin according to the van-Herk formula [4] (see *Supplement*). The margins for the MidV concept ranged from 5 to 15 mm depending on the motion and direction. A fixed 5 mm margin was used for the ITV, gating and tracking approaches mainly addressing delineation uncertainties [26].

Treatment planning

SBRT treatments with two volumetric modulated arcs were planned and optimized in Eclipse (Varian Medical System, Palo Alto, CA, USA), version 11.0.31, and doses were calculated with the analytical anisotropic algorithm with a beam energy of 6 MV in flattening filter free mode. A dose of 40.5 Gy in 3 fractions was prescribed to the 65%-isodose line around the PTV. At least 95% of the PTV had to receive the prescribed dose ($D95_{PTV}$>40.5 Gy). The dose was optimized to increase towards the tumor center, covering 95% of the TV with at least the 87%-isodose line ($D95_{TV}$>54.6 Gy). This second constraint was used to guide the high-dose to the central part of the PTV and to cover the volume of high interest with a higher dose than the PTV. The lung, heart and spinal cord doses were kept as low as possible, while for the spinal cord a strict constraint of not more than 18 Gy was used. For each

patient, a distinct treatment plan was generated for each of the six treatment concepts and dose distributions were calculated. This dose distribution is in the following referred to as the 3D dose.

Residual tracking error

Tracking latencies or mechanical restrictions of the tracking hardware can lead to an offset between the expected and actual target position. This offset, or tracking error, has to be considered during treatment planning and included into the TV either as additional margin, or as COV shift as described above. The tracking errors were evaluated for the prototype couch tracking system described by Lang et al. [10]. The measurement set-up, which was also described by Jöhl et al. [27], and evaluation of the residual motion amplitude and compensation rates are described in the *Supplement*. In summary, the motion of a robot was measured with laser triangulation (optoNCDT 1302-100, Micro Epsilon Messtechnik GmbH & Co. KG, Ortenburg, Germany) and compensated by a Protura treatment couch (CIVCO Medical Solutions, Coralville, IA, USA). The couch motion was also measured with a laser and compared to the robot's motion curve to evaluate the residual tracking error of each patient.

4D dose calculation

The 3D dose only gives an estimation of the dose, which will be received by the tumor. But the accumulated dose in the moving tumor will differ from this estimation due to the gradient, interplay and density effects. 4D dose calculations, considering these respiratory effects, were performed as earlier described in Ehrbar et al. [28]. The process is schematically drawn in Figure 1. The treatment beam was split into arc segments to create ten sub-plans, which were temporally assigned to and recalculated on the ten phase-CTs. For the amplitude-binned 4DCTs, a time-weighting of the phases was additionally introduced. For tracking, the dynamic compensation behavior of the treatment couch was considered by shifting the beam isocenter according to the tumor COV of each phase. For the reliable-couch-tracking concept, an offset between beam isocenter and tumor COV according to the residual motion was introduced at this calculation step. In the last step, the time-resolved dose was summed up according to deformable registrations between the phase CTs. This was performed with the 4D dose accumulation tool of the MIM Software (version 6.6.1). For the ITV, MidV and tracking concepts, the dose was accumulated on the initial contouring phase, and for gating on the mid-gating phase. A comparison of mid-gating and initial contouring phase as accumulation phases for gating is shown in the *Supplement*.

Dose evaluations

Target and lung dose parameters were reported. For the target, the mean dose (Dmean), the dose to 95% of the target volume (D95), the near maximum dose (D1: dose to 1% of the target) and the homogeneity index (HI) were evaluated. For HI the following definition was used:

$$HI\ (\%) = \frac{D1 - D99}{Dmean} * 100$$

These dose parameters were evaluated on the individual TV for the 3D dose, and directly on the GTV of the accumulation phase for the 4D dose. The TV used in the treatment planning were chosen for the comparison, since they are assumed to give a closer approximation of the GTV dose coverage than the PTV in presence of inhomogeneous SBRT dose distributions. Comparisons between 3D and 4D target dose parameters were performed with paired Wilcoxon sign-rank tests. Spearman correlations of the dose differences with tumor motion and with tumor size were investigated.

The 4D dose distributions were also used to report on lung dose. The mean dose (Dmean) and the relative volumes receiving more than 5 Gy and 20 Gy (V5 and V20) were evaluated for the ipsilateral lung, excluding the GTV volume. The lung parameters were normalized per patient to the values of the ITV concept. Reduction in lung dose of the different MMT were investigated. Comparisons were performed with a paired Wilcoxon sign-rank test. Spearman correlations of the relative lung doses with tumor motion and with tumor size were investigated. For all comparisons and correlations, a two-sided significance level of 5% was used (p-value < 0.05).

Results

Residual tracking error

The 20 patients showed a median motion amplitude of 3.3 mm (25- and 75-percentiles: 1.8, 6.5) and 11.6 mm (8.5, 13.2) in AP and SI direction, respectively. With couch tracking, the residual motion amplitudes were reduced to 0.8 mm (0.51, 2.21) and 2.6 mm (1.42, 3.99) in AP and SI. Detailed results are shown in the *Supplement.* The patient-individual residual motion amplitude was used to spread out the GTV volumes in the reliable couch tracking approach.

Tumor coverage

The planning constraints ($D95_{PTV}$>40.5 Gy, $D95_{TV}$>54.6 Gy, OAR constraints) were achieved for all MMT, with only minor differences in 3D dose calculations regarding tumor coverage of the TVs. The 4D dose calculations showed, that all concepts were able to cover the GTV with the prescribed PTV dose ($D95_{GTV}$>40.5 Gy). However, there were differences found when comparing the 3D dose of the TV with the 4D dose of the GTV. Figure 2 shows the 3D and 4D dose parameters. The 4D dose calculations of target Dmean and D95 gave slightly higher values than the corresponding 3D calculated parameters

for the ITV concept, gating and reliable couch tracking, while lower values were found for MidV and single-phase tracking. No significant differences in target Dmean and D95 were found for multi-phase tracking. For all MMT, except single-phase and multi-phase tracking, increases in near maximum dose (D1) were found, while HI was increased for all except gating. Overall, the 4D target dose was closest to the 3D dose for gating and multi-phase tracking. Median dose differences and percentiles are listed in Table 1 together with the correlation coefficients of the significant correlations.

For gating, the 4D dose accumulation was performed with accumulation against the mid-gating phase and also against the initial contouring phase. Here only the accumulation against mid-gating are shown, while a comparison of both accumulation phases for gating are shown in the *Supplement*.

PTV reduction and OAR sparing

When compared to the ITV concept, the single-phase tracking showed the largest reduction in PTV size (median: -40%), followed by gating, multi-phase tracking, MidV and reliable couch tracking (-25%, -25%, -23% and -16%). Analogously, a reduction in Dmean, V5 and V20 of the lung was achieved. The medians and percentiles of the absolute PTV sizes and lung doses are listed in Table 2. Figure 3 shows the PTV size and ipsilateral lung values normalized to the values of the ITV concept. The medians and percentiles of the normalized reductions are listed in Table 3 together with correlation coefficients of the significant correlations. Positive correlations to tumor size indicate a larger relative benefit for patients with smaller tumors. Negative correlations to tumor motion indicate a larger relative benefit for larger motion amplitudes. The lung benefits for MidV were mostly correlated to the tumor size, while the tracking approaches were found to be correlated primarily to tumor motion.

Discussion

The management of intrafractional lung tumor motion is important to reach dose coverage of moving tumors and to minimize dose related lung toxicities. ITV, MidV, gating and three couch tracking approaches were compared in a planning study including 4D dose calculations to evaluate their ability of lung dose sparing and tumor dose coverage. All techniques enabled a reduction in lung dose, when compared to the conservative ITV concept. Largest benefit in ipsilateral lung Dmean was shown for single-phase tracking (-29%) followed by multi-phase tracking and gating with similar reductions (-18% and -19%) and by MidV and reliable couch tracking (-15% and -12%). However, the single-phase tracking and MidV approach showed a decrease in tumor coverage (ΔD95: -3.2% and -6.3%, ΔDmean: -1.3% and -2.3%). It is worth mentioning that all techniques were able to fulfill the primary goal to deposit more than the prescribed PTV edge dose of 40.5 Gy within the GTV. With the multi-phase tracking and reliable couch tracking approaches, it was possible to consider deformations and tracking

latencies, resulting in adequate tumor coverage (ΔD95: -0.1% and +0.6%, ΔDmean: +0.1% and +0.6%), but with a lower reduction in lung dose.

The reduction in lung dose is directly related to the reduction in PTV size enabled by the MMT. Depuydt et al. [7] reported a PTV reduction of 35% (range: 16%-53%) when changing from the ITV concept to a tracking scenario. A similar reduction of 37% (10%-50%) was found here for the single-phase tracking scenarios. Guckenberger et al. [17] quantified the decrease in lung Dmean and V20, changing from ITV over to a MidV or gating approach. They showed a lower benefit for the MidV concept and a higher benefit for the gating concept compared to our work. This is mainly caused by a different adaptation of the MidV concept with dose prescription to the accumulated GTV and by the smaller gating duty cycle of only 12.5% employed by them, compared to our 30%.

The 4D dose calculation showed a slight increase in target Dmean and D95 for the ITV, gating and reliable couch tracking approaches when compared to the 3D dose. This can be explained by the use of a planning TV which is larger than the GTV. Treatment plans for SBRT also show a dose increase within the TV with the lowest TV doses at the edges of the volume. The moving GTV itself is not always occupying the lower dose regions and therefore accumulates a higher dose than estimated by the TV. A more prominent decrease in target Dmean and D95 was shown for MidV and single-phase tracking. For the MidV concept, this can be explained by the nature of the concept itself, where 90% of the patients were expected to receive 95% of the prescription dose. The decrease for the single-phase tracking approach is of a different nature. The TV was restricted to the GTV of one phase. Deformations and rotations of the tumor throughout the respiratory phases were neglected at the planning stage. These changes of the tumor volume over the respiratory cycle were considered in the multi-phase tracking approach, which was able to fully account for these effects.

The changes seen in target D1 can be attributed mainly to gradient and interplay effects. The single-phase and multi-phase tracking show only a spread out in D1, but no significant increase, whereas for the other concepts, the D1 was increased.

Guckenberger et al. [29] showed that the biologically effective dose (BED) to the CTV, estimated with the 4D calculation, did better fit the tumor control probability curve than the 3D doses at the PTV margin. Further studies showed that the PTV maximal dose is a better prognostic factor for tumor control than the PTV enclosing dose [30, 31] and that dose prescription to the GTV median dose rather than the PTV enclosing provides a more robust method for treating lung lesions with SBRT [32]. An evaluation of the 4D dose to the CTV, in our case GTV, is therefore recommendable for the comparison of dynamic treatment delivery techniques. The fractionation scheme and prescription dose of the present study resulted in a biological effective dose (BED) of 192 Gy BED to the PTV maximal dose. Also the estimated 4D maximal dose to the GTV ranged from 182 to 200 Gy BED over all techniques and

patients. In this dose region, the techniques are expected to result in similar tumor control probabilities. This conclusion might change with a lower dose prescription.

End-of-inhale was chosen as gating window and the dose for the gating plan was accumulated to the middle-most phase of the gating window (mid-gating). The decrease in lung dose for the gating approach is therefore not only caused by the decrease in PTV size, but also by the increase in lung volume at end-of-inhale. For all the other MMT, the initial contouring phase was used for dose accumulation to minimize registration errors in the dose accumulation process and to evaluate on the same contours. However, the gating plans would only be applied to the patient with the tumor in the gating window. Therefore, the accumulation to the mid-gating phase represents best the accumulated tumor dose. However, choosing the mid-gating phase as accumulation phase might give an advantage in the tumor coverage for gating, since all phases are close to the accumulation phase and therefore might be less influenced by registration errors.

Another way to evaluate accumulated tumor dose are direct dose measurements in a motion phantom. This was performed in one of our former studies [33] with an anthropomorphic, deformable thorax phantom. For gating and tracking, superior dose coverage and better lung sparing was shown than for ITV and MidV concepts. For tracking, a single-phase planning approach was used, since the tumor was a rigid sphere, and resulted in similar dose coverage to the gating approach.

Limitations of a couch tracking system were considered in the reliable couch tracking approach by including the residual motion amplitude in the TV delineation. The latency and mechanical limitations cause the couch to lag behind the tumor position. The residual motion is thus largest at the steep motion gradients. The Protura couch was not initially designed for dynamic motion compensation and its maximal speed is restricted to 16 mm/s. Patients with tumor motion faster than this will be insufficiently compensated. Couch tracking systems with higher compensation speed, as for example the PerfectPitch couch (Varian), would better compensate the tumor motion and therefore reduce the residual motion.

Including the residual tracking errors into the TV definition resulted in increased target volumes, and therefore less sparing of the lung. In the presented reliable couch tracking approach, the treatment volume was increased to cover the full extent of residual motion. However, the tumor motion is only offset during a small amount of the overall treatment time. This indicates that the used approach is overprotective and it is cumbersome to conduct a patient-individual evaluation and TV adaption. Alternatively, a combined tracking-MidV approach could be considered, including the residual motion and tumor deformation into the PTV margin as statistical uncertainties. Patient individual estimations of the residual motion can be made based on the respiration during CT acquisition in combination with a performance model of the treatment couch [27].

A laser system was used as motion feedback for the couch tracking. It is only capable of measuring surface motions. To detect internal motion, fluoroscopy imaging of gold markers or detection of implanted electromagnetic beacons can be used [34, 35, 36]. These systems, however, have a longer latency time than the employed laser system. An increased latency would also increase the residual motion. This effect can be mitigated by applying prediction filters [37] or regularly updated external-internal motion correlations [38].

The 4DCT images do not fully reduce image artifacts, caused by the tumor motion [39, 40, 41]. These artifacts affect the tumor delineation, TV definition and 4D dose calculation. It is therefore questionable to what extent the differences seen in the GTVs of the ten phases are caused by image artifacts or by real deformations of the tumor, and which tracking concept (single- or multi-phase) is more appropriate.

For the comparison of different dynamic treatment techniques, the respiration dynamics have to be considered to evaluate the ability of each technique to cover the tumor with dose or to spare healthy tissue. The effects of the respiratory behavior of the tumor were therefore considered in 4D dose calculations. Further uncertainties as set-up and delineation errors or changes in respiratory cycle were neglected in the calculations, however, they were expected to affect the techniques equally.

The 4D dose calculations were based on several assumptions. First, a regular respiratory period (3.4 s for phase-based and 5.1-9.18 s for amplitude-based reconstructed 4DCTs) was assumed to facilitate the temporal assignment of the arc segments to the phases. Second, fractionation effects were neglected in the calculation, which could mitigate the impact of the interplay effect and third, the assignment of the arc segments was done with regarding only one starting phase per patient. Rao et al. [42] showed that the 4D-to-3D dose differences for lung SBRT using an ITV concept are only to a minor extent caused by the interplay effect and also the choice of the respiratory starting phase for the assignment of the arc segments did not change the difference substantially. This would mean that the 4D calculations were assumed to be insensitive against the length of the respiratory period and the assignment starting phase. We did evaluate the dependence on the respiratory starting phase for one patient (see supplement). For the gating and tracking treatments, no relevant dependence was found (spread < 0.4 Gy), but for ITV and MidV concept, the Dmean and D1 showed a spread of 0.8 Gy and 1.7 Gy, respectively, for different starting phases. By including 20 patients to the study but only one starting phase, we considered to sufficiently cover the interplay variation by the patient sample size. To better judge the impact of the interplay effect for different motion-management techniques, elaborate investigations beyond the scope of this study should be performed.

Also the accuracy of the deformable image registration directly influences the accuracy of the 4D dose calculation. Kadoya et al. [43] showed for the MIM software a mean 3D registration error of 3.29 mm between end-of-inhale to end-of-exhale CT phases and that the likelihood of registration error

increased with initial displacement. In the present work, the registrations of the 4DCT phases were performed against a respiratory phase in the middle of the breathing cycle. This reduced the maximal displacement to be registered by the deformable image registration algorithm and thereby minimized the registration error. On the other hand, the middle phase is more prone to image artifacts caused by tumor motion which restrict the accuracy of the 4D dose calculations.

This study is valid for dynamic treatment-couch tracking at a conventional linear accelerator, but might also be valid for other tracking types to some extent. As planning strategy, a single- or multi-phase tracking approach is often used, but the delivery technique might differ between tracking types and change the results.

Conclusion

The use of the ITV concept for lung SBRT ensures excellent tumor coverage, but exposes the healthy lung tissue to an unnecessarily high amount of radiation. The treatment volume and lung doses can easily be reduced by employing other motion-management techniques as MidV, gating or couch tracking concepts. However, 4D dose calculations should be employed to evaluate the tumor dose coverage for these techniques. Tumor deformation and system latencies should be considered in the treatment planning approaches for tumor tracking to fully ensure target coverage.

Acknowledgements

This work was supported by the grant CR32I3_153491 of the Swiss National Science Foundation.

Table 1: GTV coverage: Median values of the differences between 4D and 3D dose calculations (4D – 3D value) are listed together with 25%- and 75%-percentiles. Correlation coefficients of the differences with tumor motion amplitude or tumor size are given for the significant correlations ($p<0.05$).

	ΔDmean (%)				ΔD95 (%)			
	Median	(prctiles)	R motion	R size	Median	(prctiles)	R motion	R size
ITV Concept	0.8	(0.4, 1.2)	-	-	1.1	(0.2, 1.4)	-	-
MidV Concept	-2.3	(-3.0, -1.7)	-	-	-6.3	(-8.5. -4.9)	-	-
Gating	0.4	(0.4, 0.9)	-	-	0.7	(0.4, 1.1)	-	-
Single-phase Tracking	-1.3	(-2.0, -0.9)	-	0.52	-3.2	(-5.1, -2.4)	-0.49	-
Multi-phase Tracking	0.1	(-0.3, 0.4)	-	0.67	-0.1	(-0.6, 0.6)	-	0.58
Reliable Couch Tracking	0.6	(0.0, 0.8)	-	0.54	0.6	(0.0, 1.2)	-	0.59
	ΔD1 (%)				ΔHI (pp)			
	Median	(prctiles)	R motion	R size	Median	(prctiles)	R motion	R size
ITV Concept	2.4	(1.1, 3.2)	-	-0.47	2.5	(0.2, 3.7)	-	-
MidV Concept	0.9	(0.0, 1.6)	-	-	10.2	(7.4, 12.1)	-	-
Gating	0.2	(0.1, 0.6)	-	-	0.1	(-0.5, 0.6)	-	-
Single-phase Tracking	-0.3	(-0.7, 0.4)	-	0.78	5.3	(3.0, 7.8)	0.56	-
Multi-phase Tracking	0.0	(-0.2, 0.4)	-	0.58	1.0	(0.0, 3.1)	0.61	
Reliable Couch Tracking	0.6	(0.0, 1.3)	-	0.55	0.5	(0.2, 1.6)	0.49	-

prctiles: (25-percentile, 75-percentile), R: Spearman rank correlation coefficient, Δ: 4D minus 3D value

Table 2: Absolute PTV size and ipsilateral lung dose parameters: Median values of PTV size, lung Dmean, V5 and V20 are listed together with the 25%- and 75%-percentiles.

	PTV size (cc)		Dmean (Gy)		V5 (%)		V20 (%)	
	Median	(prctiles)	Median	(prctiles)	Median	(prctiles)	Median	(prctiles)
ITV Concept	32	(23, 72)	4.8	(3.2, 6.6)	27.8	(19.7, 35.8)	6.4	(3.2, 8.0)
MidV Concept	24	(17, 59)	4.3	(2.6, 5.7)	24.2	(15.7, 32.3)	4.8	(2.5, 6.5)
Gating	23	(17, 56)	3.8	(2.7, 5.0)	22.2	(15.5, 26.9)	4.3	(2.3, 5.8)
Single-phase Tracking	19	(14, 44)	3.3	(2.4, 4.7)	20.1	(13.9, 26.4)	3.7	(2.0, 4.8)
Multi-phase Tracking	23	(16, 51)	3.6	(2.8, 5.1)	21.4	(14.8, 28.2)	4.5	(2.4, 5.9)
Reliable Couch Tracking	26	(19, 60)	4.1	(2.9, 5.6)	23.2	(15.8, 32.4)	5.3	(2.7, 6.8)

prctiles: (25-percentile, 75-percentile)

Table 3: Normalized PTV size and ipsilateral lung dose parameters: Median reduction of the parameters (PTV size, lung Dmean, V5 and V20), normalized by the values of the ITV concept plan, are listed together with the 25%- and 75%-percentiles. Correlation coefficients are given for the significant correlations ($p<0.05$) of the reduction with tumor motion or tumor size.

	Δ relative PTV size (%)				Δ relative Dmean (%)			
	Median	(prctiles)	R motion	R size	Median	(prctiles)	R motion	R size
MidV Concept	-23	(-29, -18)	-	0.45	-15	(-20, -11)	-	-
Gating	-25	(-30, -18)	-0.57	-	-19	(-27, -16)	-0.48	-
Single-phase Tracking	-40	(-42, -33)	-0.69	-	-29	(-32, -23)	-0.68	-
Multi-phase Tracking	-25	(-33, -21)	-0.81	-	-18	(-24, -14)	-0.79	
Reliable Couch Tracking	-16	(-22, -11)	-0.55	-	-12	(-17, -8)	-0.47	0.47
	Δ relative V5 (%)				Δ relative V20 (%)			
	Median	(prctiles)	R motion	R size	Median	(prctiles)	R motion	R size
MidV Concept	-11	(-18, -7)	-	0.58	-22	(-30, -17)	-	-
Gating	-18	(-25, -12)	-	0.51	-26	(-34, -20)	-	-
Single-phase Tracking	-28	(-33, -18)	-0.47	0.58	-38	(-41, -34)	-0.47	-
Multi-phase Tracking	-19	(-26, -10)	-0.57	-	-26	(-30, -18)	-0.67	-
Reliable Couch Tracking	-14	(-21, -7)	-0.45	-0.53	-17	(-20, -8)	-0.47	-

prctiles: (25-percentile, 75-percentile), R: Spearman rank correlation coefficient, Δ: Normalized difference to ITV concept

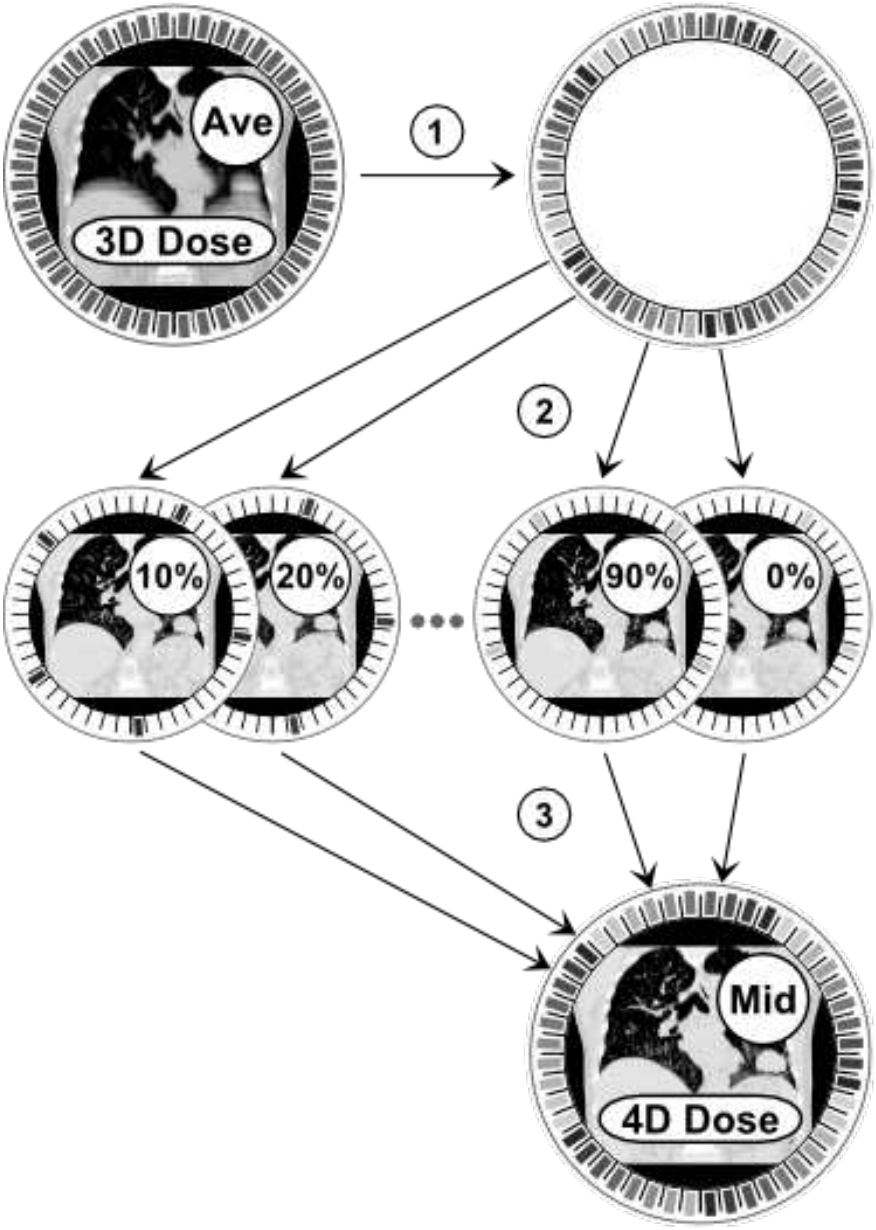


Figure 1: 4D dose calculation: (Top left) The initial treatment plan with its 3D dose distribution on the planning CT (here the average CT). (1) Separation of the treatment plan into arc segments assigned to the respiratory phases (different gray levels). (2) Recalculation of time-resolved sub-plans on the phase CTs, resulting in a time-resolved dose. (3) Dose summation on a mid-respiration phase according to deformation fields resulting in a 4D dose distribution.

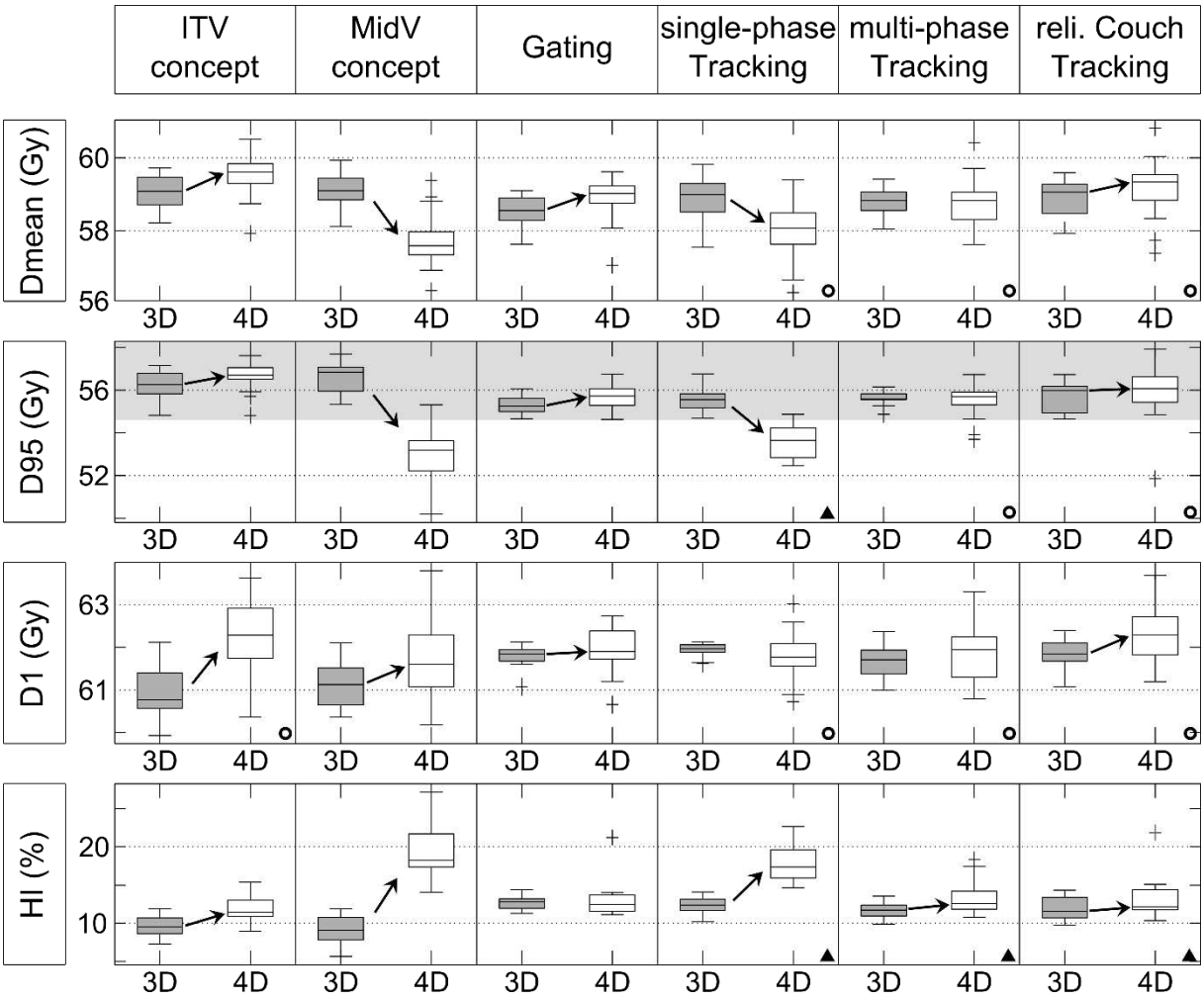


Figure 2: Results of the 3D and 4D dose calculations: The Dmean, D95, D1 and HI of the TV from the original treatment plans (3D: gray boxplots) and the accumulated GTV doses (4D: white boxplots). Planning constraints for D95 are highlighted in light gray. Arrows indicate significant differences between 3D and 4D dose. Triangles and circles indicate significant correlations of these differences with tumor motion amplitude or tumor size, respectively.

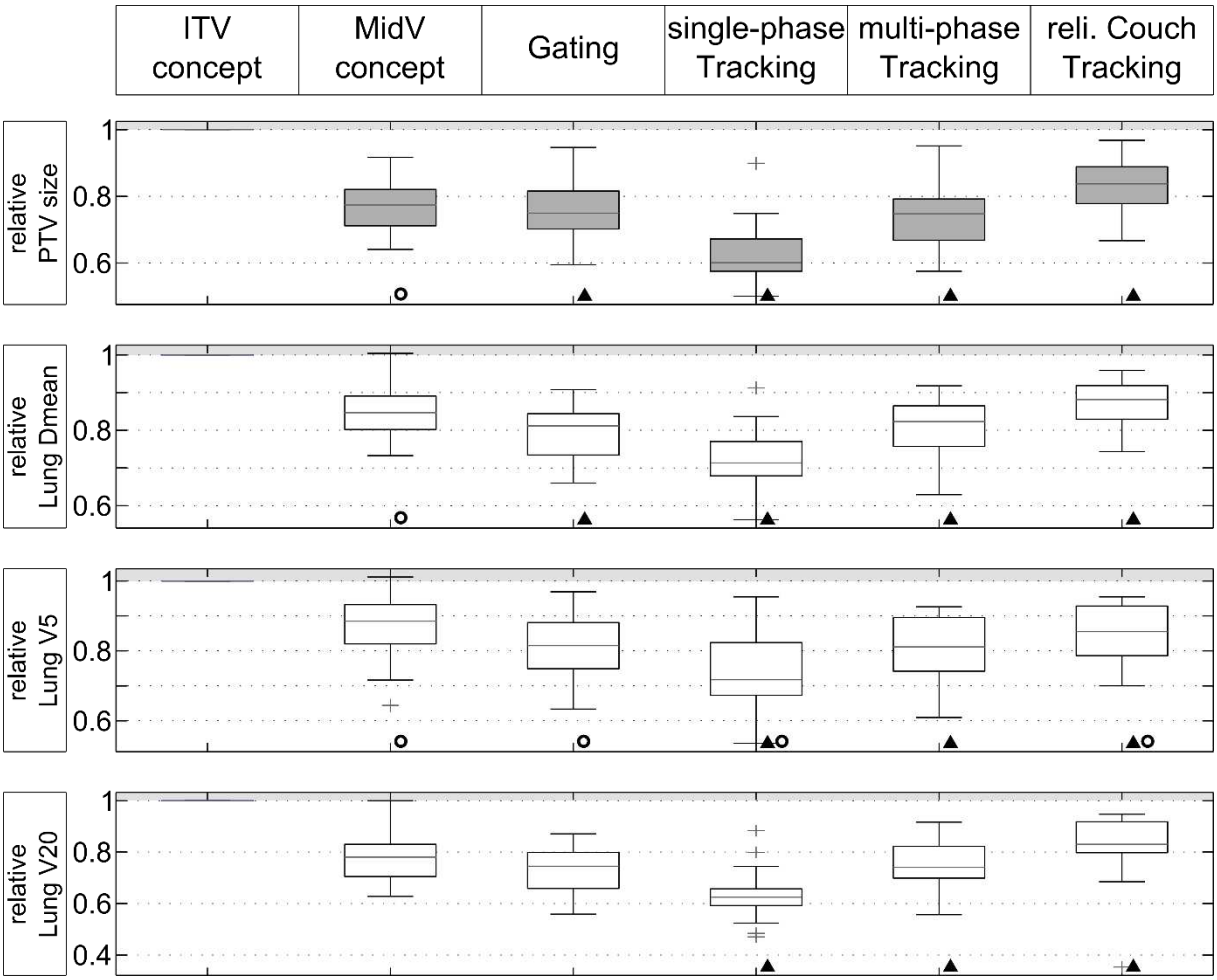


Figure 3: Normalized PTV size and 4D dose parameters of the ipsilateral lung: The parameters were normalized with the values of the ITV concept plan for each patient. Triangles and circles indicate significant correlations of the normalized parameters with tumor motion amplitude (all positively correlated) or tumor size (all negatively correlated), respectively.

References

[1] Erridge SC, Seppenwoolde Y, Muller SH et al. Portal imaging to assess set-up errors, tumor motion and tumor shrinkage during conformal radiotherapy of non-small cell lung cancer. Radiotherapy and Oncology 2003;66(1):75-85.

[2] Nyman J, Hallqvist A, Lund J-A et al. SPACE – A randomized study of SBRT vs conventional fractionated radiotherapy in medically inoperable stage I NSCLC. Radiotherapy and Oncology 2016; 121(1): 1-8.

[3] Wambersie A, Landgerg T. ICRU Report 62: Prescribing, Recording and Reporting Photon Beam Therapy. ICRU Publ Bethesda MD. 1999.

[4] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. International Journal of Radiation Oncology* Biology* Physics 2000;47(4):1121-35.

[5] Ohara D, Okumura T, Akisada M et al. Irradiation synchronized with respiration gate. International Journal of Radiation Oncology* Biology* Physics 1989;17(4):853-57.

[6] Keall P, Kini V, Vedam S, Mohan R. Potential radiotherapy improvements with respiratory gating. Australasian Physics & Engineering Sciences in Medicine 2002;25(1):1-6.

[7] Depuydt T, Poels K, Verellen D et al. Treating patients with real-time tumor tracking using the Vero gimbaled linac system: implementation and first review. Radiotherapy and Oncology 2014;112(3):343-51.

[8] Adler J, Murphy M, Chang SD, Hancock S. External respiratory motion analysis. Neurosurgery 1999;44(6):1299-307.

[9] Neicu T, Shirato H, Seppenwoolde Y, Jiang S. Synchronizedmoving aperture radiation therapy (SMART): Average tumour trajectory for lung patients. Physics in Medicine and Biology 2003;48(5):587-98.

[10] Lang S, Zeimetz J, Ochsner G, Daners MS, Riesterer O, Klöck S. Development and evaluation of a prototype tracking system using the treatment couch. Medical physics 2014;41(2):021720.

[11] Buzurovic I, Huang K, Yu Y, Podder T. A robotic approach to 4D real-time tumor tracking for radiotherapy. Physics in Medicine and Biology 2011;56(5):1299-318.

[12] Haas OCL, Skworcow P, Paluszczyszyn D, Sahih A, Ruta M, Mills J. Couch-based motion compensation: Modelling, simulation and real-time experiments. Physics in Medicine and Biology 2012;57(18):5787-807.

[13] Wilbert J, Meyer J, Baier K et al. Tumor tracking and motion compensation with an adaptive tumor tracking system (ATTS): System description and prototype testing. Medical Physics 2008;35(9):3911-21.

[14] Guckenberger M, Wilbert J, Krieger T et al. Four-dimensional treatment planning for stereotactic body radiotherapy. International Journal of Radiation Oncology* Biology* Physics 2007;69(1):276-85.

[15] Admiraal MA, Schuring D, Hurkmans CW. Dose calculations accounting for breathing motion in stereotactic lung radiotherapy based on 4D-CT and the internal target volume. Radiotherapy and Oncology 2008;86(1):55-60.

[16] Mexner V, Wolthaus JW, van Herk M, Damen EM, Sonke J-J. Effects of respiration-induced density variations on dose distributions in radiotherapy of lung cancer. International Journal of Radiation Oncology* Biology* Physics 2009;74(4):1266-75.

[17] Guckenberger M, Kavanagh A, Webb S, Brada M. A novel respiratory motion compensation strategy combining gated beam delivery. Radiotherapy and Oncology 2011;98(3):317-22.

[18] Menten MJ, Fast MF, Nill S, Kamerling CP, McDonald F, Oelfke U. Lung stereotactic body radiotherapy with an MR-linac-Quantifying the impact of the magnetic field and real-time tumor tracking. Radiotherapy and Oncology 2016;119(3):461-66.

[19] Ong C, Verbakel WF, Cuijpers JP, Slotman BJ, Senan S. Dosimetric impact of interplay effect on RapidArc lung stereotactic treatment delivery. International Journal of Radiation Oncology* Biology* Physics 2011;79(1):305-11.

[20] Chin E, Loewen S, Nichol A, Otto K. 4D VMAT, gated VMAT, and 3D VMAT for stereotactic body radiation therapy in lung. Physics in medicine and biology 2013;58(4):749.

[21] Chin E, Otto K. Investigation of a novel algorithm for true 4D-VMAT planning with comparison to tracked, gated and static delivery. Medical physics 2011;38(5):2698-707.

[22] Falk M, Pommer T, Keall P et al. Motion management during IMAT treatment of mobile lung tumors—A comparison of MLC tracking and gated delivery. Medical physics 2014;41(10)101707.

[23] Menten MJ, Guckenberger M, Herrmann C et al. Comparison of a multileaf collimator tracking system and a robotic treatment couch tracking system for organ motion compensation during radiotherapy. Medical physics 2012;39(11):7032-41.

[24] Wolthaus JW, Sonke J-J, van Herk M et al.Comparison of different strategies to use four-dimensional computed tomography in treatment planning for lung cancer patients. International Journal of Radiation Oncology* Biology* Physics 2008;70(4):1229-38.

[25] Lang S, Shrestha B, Graydon S et al. Clinical application of flattening filter free beams for extracranial stereotactic radiotherapy. Radiotherapy and Oncology 2013;106(2):255-59.

[26] Sonke J-J, Rossi M, Wolthaus J, van Herk M, Damen E, Belderbos J. Frameless stereotactic body radiotherapy for lung cancer using four-dimensional cone beam CT guidance. International Journal of Radiation Oncology* Biology* Physics 2009;74(2):567-74.

[27] Jöhl A, Lang S, Ehrbar S et al. Modeling and performance evaluation of a robotic treatment couch for tumor tracking. Biomedical Engineering/Biomedizinische Technik 2016;61(5):557-66.

[28] Ehrbar S, Lang S, Stieb S et al. Three-dimensional versus four-dimensional dose calculation for volumetric modulated arc therapy of hypofractionated treatments. Zeitschrift für Medizinische Physik 2016;26(1):45-53.

[29] Guckenberger M, Wulf J, Mueller G et al. Dose-response relationship for image-guided stereotactic body radiotherapy of pulmonary tumors: relevance of 4D dose calculation. International Journal of Radiation Oncology* Biology* Physics 2009;74(1):47-54.

[30] Klement RJ, Allgäuer M, Andratschke N et al. Bayesian Cure Rate Modeling of Local Tumor Control: Evaluation in Stereotactic Body Radiation Therapy for Pulmonary Metastases. International Journal of Radiation Oncology* Biology* Physics 2016;94(4):841-49.

[31] Guckenberger M, Klement RJ, Allgäuer M et al. Local tumor control probability modeling of primary and secondary lung tumors in stereotactic body radiotherapy. Radiotherapy and Oncology 2016;118(3):485-91.

[32] Lacornerie T, Lisbona A, Mirabel X, Lartigau E, Reynaert N. GTV-based prescription in SBRT for lung lesions using advanced dose calculation algorithms. Radiation Oncology 2014;9(1):223.

[33] Ehrbar S, Perrin R, Peroni M et al. Respiratory motion-management in stereotactic body radiation therapy for lung cancer - A dosimetric comparison in an anthropomorphic lung phantom (LuCa). Radiotherapy and Oncology 2016;121(2):328-34.

[34] Hoisak J, Katharina E, Tirona R, Cheung P, Pignol J-P. Prediction of lung tumour position based on spirometry and on abdominal displacement: Accuracy and reproducibility. Radiotherapy and Oncology 2006;78(3):339-46.

[35] Keall PJ, Colvill E, O'Brien R et al. The first clinical implementation of electromagnetic transponder-guided MLC tracking. Medical physics 2014;41(2):020702.

[36] Shimizu S, Shirato H, Ogura S et al. Detection of lung tumor movement in real-time tumor-tracking radiotherapy. International Journal of Radiation Oncology* Biology* Physics 2001;51(2):304-10.

[37] Ernst F, Dürichen R, Schlaefer A, Schweikard A. Evaluating and comparing algorithms for respiratory motion prediction. Physics in medicine and biology 2013;58(11):3911.

[38] Schweikard A, Shiomi H, Adler J. Respiration tracking in radiosurgery. Medical Physics 2004;31(10):2738-41.

[39] Nakamura M, Narita Y, Sawada A et al. Impact of motion velocity on four-dimensional target volumes: A phantom study. Medical physics 2009;36(5):1610-17.

[40] Watkins WT, Li R, Lewis J et al. Patient-specific motion artifacts in 4DCT. Medical physics 2010;37(6):2855-61.

[41] Persson GF, Nygaard DE, Brink C et al. Deviations in delineated GTV caused by artefacts in 4DCT. Radiotherapy and Oncology 2010;96(1):61-66.

[42] Rao M, Wu J, Cao D, Wong T, Mehta V, Shepard D, Ye J. Dosimetric impact of breathing motion in lung stereotactic body radiotherapy treatment using image-modulated radiotherapy and volumetric modulated arc therapy. International Journal of Radiation Oncology* Biology* Physics 2012;83(2):e251-56.

[43] Kadoya N, Nakajima Y, Saito M et al. Multi-institutional validation study of commercially available deformable image registration software for thoracic images. International Journal of Radiation Oncology* Biology* Physics 2016;96(2):422-31.

Supplement
ITV, Mid-Ventilation, Gating or Couch Tracking – A comparison of respiratory motion-management techniques based on 4D dose calculations

S Ehrbar, A Jöhl, A Tartas, LS Stark, O Riesterer, S Klöck, M Guckenberger, S Tanadini-Lang

Contents

1 Patient characteristics **1**

2 Margin calculation **1**

3 Residual tracking error **2**
3.1 Methods 2
3.2 Results 3

4 Dependence on accumulation phase for gating **4**

5 Dependence on starting phase for 4D dose calculation **5**

1 Patient characteristics

Out of the twenty patients, eight were treated for primary disease (four non-small and four small cell lung cancer) and twelve had metastatic lesions. Most selected lesions were located in the lower lung lobe, since patients with pronounced tumor motion were preferred in the selection. Respiratory periods ranged from 1.8 to 5.8 s, gross tumor volumes (GTV) from 0.6 to 79.3 cc and 3D tumor motion amplitudes from 4.9 to 27.8 mm (see Table 1 for details).

2 Margin calculation

The planning target volumes (PTV) were derived by adding a margin to the target volume (TV). The margin was calculated using systematic (Σ) and random (σ) positioning errors of the tumor according to an adapted version of the van-Herk formula [1]:

$$M = 2.5 * \sqrt{\Sigma_{Setup}^2 + \Sigma_{BL}^2 + \Sigma_{Del}^2 + \Sigma_{PD}^2} + \beta * \sqrt{\sigma_{Setup}^2 + \sigma_{BL}^2 + \sigma_{TM}^2 + \sigma_P^2} - \beta * \sigma_P$$

For a margin to ensure coverage of the target volume with 95% of the dose in 90% of the cases, a β-value of 1.64 was used. A σ_P of 6.4 mm was used (according to [2]) to account for the beam penumbra in lung tissue. Setup errors (Σ_{Setup}, σ_{Setup}) were assumed to be compensated using CB-CT imaging and therefore set to zero. For baseline shifts over the course of treatment, values from Lang et al. [3] were used (Σ_{BL}=0.99 mm, σ_{BL}=1.08 mm). For delineation uncertainties, a value of Σ_{Del}=1.70 mm was used. Only for the MidV margin, the tumor motion was taken into the calculation as random

Supplementary Table 1: Patient Characteristics: Tumor type and localization, respiratory period from RPM curves, GTV size and tumor motion amplitude from 4D-CT.

					Tumor motion (mm)			
Patient	Type	Localization	Period (s)	GTV size (cc)	LR	AP	SI	3D
1	NSCLC	l LL	2.7	79.3	1.7	1.8	4.2	4.9
2	NSCLC	l LL	3.8	6.0	1.9	7.4	17.7	19.2
3	Metastasis	r LL	4.8	3.0	3.0	2.7	6.3	7.5
4	Metastasis	l LL	3.0	0.6	0.8	1.8	5.1	5.5
5	Metastasis	l LL	3.0	37.7	4.3	2.9	21.8	22.4
6	NSCLC	l LL	3.3	17.3	1.0	5.3	11.8	13.0
7	SCLC	r UL	4.8	4.4	3.9	3.6	7.2	9.0
8	Metastasis	l LL	4.2	7.2	0.8	4.6	11.0	11.9
9	Metastasis	r ML	4.6	7.8	2.7	1.9	9.1	9.7
10	Metastasis	l LL	4.8	3.7	2.8	3.0	14.1	14.7
11	NSCLC	r LL	5.8	16.4	6.2	8.9	25.6	27.8
12	Metastasis	l LL	4.0	3.9	3.5	7.0	17.4	19.0
13	Metastasis	l LL	4.5	2.3	1.6	0.7	12.8	13.0
14	Metastasis	l LL	4.2	15.5	4.2	7.3	11.9	14.5
15	SCLC	l LL	3.8	25.2	1.8	7.2	11.5	13.6
16	SCLC	r ML	3.9	20.8	0.9	1.2	11.9	12.0
17	Metastasis	r UL	5.7	2.2	1.1	4.4	12.1	12.9
18	SCLC	r LL	3.1	3.8	1.9	6.4	9.8	11.9
19	Metastasis	l LL	1.7	9.1	1.6	1.2	6.6	6.9
20	Metastasis	l LL	3.2	1.8	1.1	1.1	9.0	9.1
Median			3.9	6.6	1.9	3.3	11.6	12.5
25-prctile			3.1	3.5	1.1	1.8	8.5	9.1
75-prctile			4.6	16.6	3.1	6.5	13.2	14.6

(N)SCLC: (Non) Small Cell Lung Cancer, l: left, r: right, LL/ML/UL: Lower/Middle/Upper Lobe, prctile: percentile

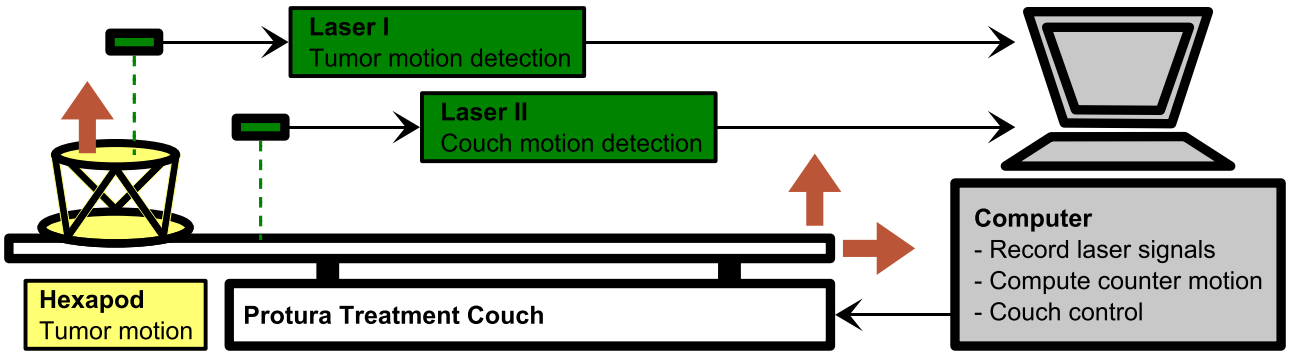
positioning error (σ_{TM}). The standard deviation of the motion (= random error) was approximated to be 1/3 of the motion amplitude in each direction. Additionally, a fraction of 1/10 of the motion was used as systematic error (Σ_{PD}) to account for the binned data representation of the 4D-CT phases which could cause an error in the phase determination (PD) of the MidV phase. The MidV margins ranged from 5 to 15 mm depending on the motion. A 5-mm margin for no motion was used for the ITV, gating and tracking approaches.

3 Residual tracking error

3.1 Methods

For the estimation of the residual motion during tracking, internal tumor motion curves of the patients are required. This internal tumor motion was simulated by scaling the external RPM signal during 4D-CT acquisition with the internal tumor motion amplitude retrieved from the 4D-CT. The signal was cut to 100 seconds and scaled, such that the 5%-to-95%-percentile range of the motion signal equaled the internal tumor motion amplitude from the 4D-CT. This patient specific motion was then simulated with the robotic hexapod H-840.5PD (Physik Instrumente GmbH & Co. KG, Karlsruhe/Palmbach,

Germany). The motion of the hexapod was measured via laser triangulation using the displacement sensor optoNCDT 1302-100 (Micro Epsilon Messtechnik GmbH & Co. KG, Ortenburg, Germany), which was attached to the treatment couch. This laser system has a latency of 1 ms, a measuring rate of 500 Hz and a resolution of 50 μm. The laser signal was used as input signal for the tracking system. A Simulink interface (The MathWorks, Inc., Natick, MA, USA) managed the conversion of the input signal to couch control signals. No prediction filter was applied to the signal. The Protura treatment couch (CIVCO Medical Solutions, Coralville, IA, USA) was employed to compensate the motion. The speed of the robotic couch system was limited to 16 mm/s. The couch and laser system combined had a total system latency of 57 ms [4]. The couch was set to compensate the motion in superior-inferior (SI) and anterior-posterior (AP) direction. The couch motion was monitored with a second laser system attached to the immobile floor. The measurement setup is shown in Figure 1.



Supplementary Figure 1: Measurement setup: The robotic Hexapod simulates the patient's tumor motion, which is monitored by Laser I, attached to the table. The detected motion is converted to couch control signals to compensate this motion in either AP or SI direction. The motion of the treatment couch is monitored by Laser II, attached to the floor. The motion signals from Laser I and II can be compared to calculate the residual tumor motion.

The two laser signals were compared against each other. In an ideal tracking scenario, the sum of both signals would always equal zero. In a non-ideal tracking scenario, the sum corresponds to the residual offset between plan isocenter and beam isocenter. The 5%-to-95%-percentile difference of the residual motion signal was calculated and defined as the residual motion range, which was used for the definition of the tiTV. The tracking performance was additionally evaluated by calculating an amplitude compensation rate (ACR) using the ratio of the residual motion amplitude (A_{res}) and the initial tumor motion amplitude (A):

$$ACR(\%) = (1 - \frac{A_{res}}{A}) * 100$$

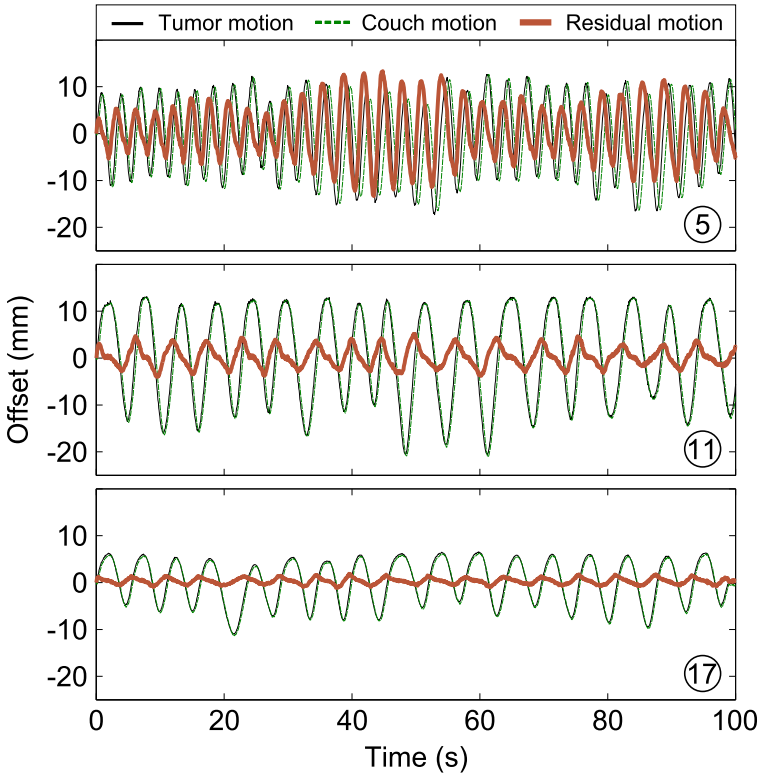
A value of 100% means full compensation of the tumor motion by the treatment couch, a value of 0% means no compensation at all. The root-mean-square error ($RMSE$) of the tumor motion signal and the residual motion signal were calculated. The tracking performance was also evaluated by calculating an error compensation rate (ECR) using the ration of the residual RMSE ($RMSE_{res}$) and the initial RMSE ($RMSE_{init}$):

$$ECR(\%) = (1 - \frac{RMSE_{res}}{RMSE_{init}}) * 100$$

The ECR is less influenced by short-duration spikes, which occur at the steep motion gradients, than the ACR and therefore gives a more holistic performance evaluation. To fully ensure tumor coverage in the tracking-ITV approach, the ACR was used for contouring.

3.2 Results

The laser signals of three measurements for the determination of the residual motion amplitude are shown in Figure 2. These represent the measurements with the worst, the median and the best



Supplementary Figure 2: Residual motion in SI: The tumor motion (black), simulated with the Hexapod, and the negative couch counter-motion (dark gray) were measured simultaneously with two triangulation laser sensors. The difference of these signal is the residual tumor motion (thick bright gray). The results with the worst, median and best tracking performance are shown (patient 5, 11 and 17).

amplitude compensation rate. The residual motion is largest in between inhale and exhale, where the steepest motion gradients occur. The evaluated residual motion amplitudes, RMSE and compensation rates are shown in Table S2. Over all patients, median ACR in AP and SI direction of 71% and 77%, and ECR of 89% and 95%, respectively, were found.

4 Dependence on accumulation phase for gating

In the dose accumulation process of the 4D dose calculation, an accumulation phase has to be chosen. The deformation fields between the phases and this accumulation phase are used to warp the dose. In the manuscript, the mid-gating phase in the middle of the gating window was chosen as accumulation phase, while for the other methods, the initial countouring phase inbetween inhale and exhale was used. Choosing the accumulation phase in the middle of all relevant phases minimizes the registration error, which increases with increased distance of the lesions inbetween the phases. For gating, all relevant phases (i.e. phases with applied dose) are close to the accumulation phase. This gives an advantage to gating, since the other methods have also some relevant phases further away from the accumulation phase. The alternative would be to use the same accumulation phase for gating as for the other methods. However, this poses a disadvantage to the accumulated gating phase, since all relevent phases are at end of inhale and threrfore maximally influenced by the registration error. Both strategies for dose accumulation are compared in Figure 3, with 4D dose accumulation against the initial contouring phase ($4D_{Mid}$) and against the mid-gating phase at end of inhale ($4D_{EOI}$). The tumor coverage (D95) was slightly inferior for the $4D_{Mid}$ dose compared to the 3D and $4D_{EOI}$ dose. Especially, there were two outliers with drastically lower Dmean and D95 in the $4D_{Mid}$ accumulation. No significant difference in the maximal dose (D1) was found.

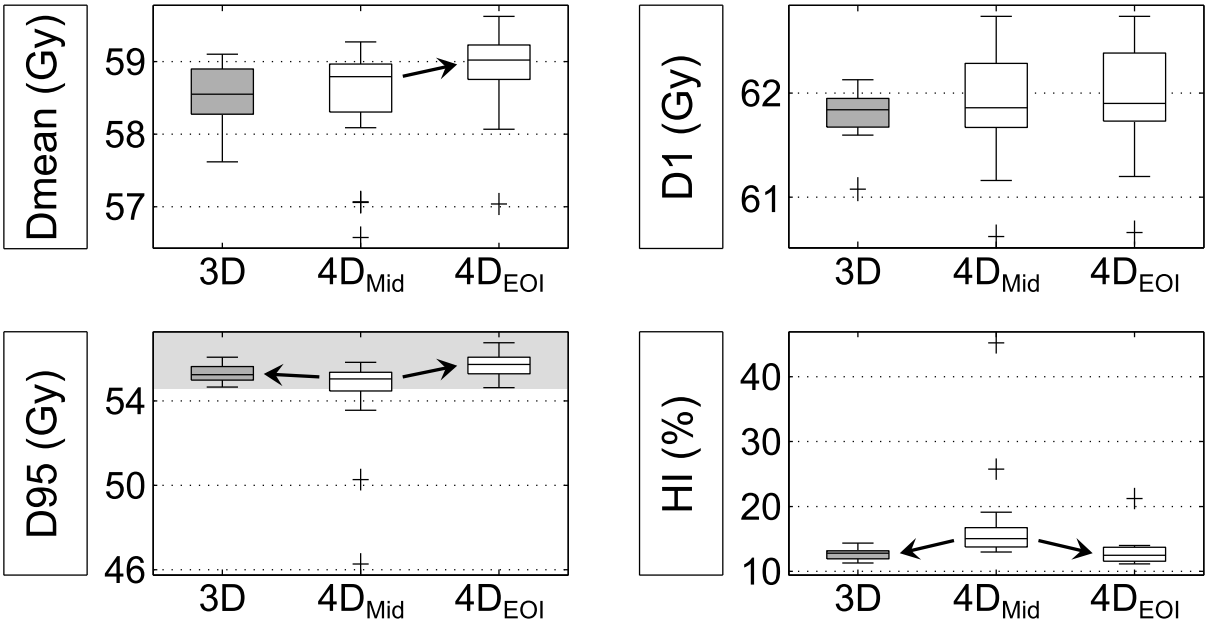
Supplementary Table 2: Tracking performance: Residual motion amplitude, root mean square error (RMSE), amplitude and error compensation rates (ACR and ECR) for SI and AP motion compensation.

	Res. Amplitude (mm)		ACR (%)		Initial RMSE (mm)		Res. RMSE (mm)		ECR (%)	
Patient	AP	SI	AP	SI	AP	SI	AP	SI	AP	SI
1	0.50	0.96	72	77	0.44	2.38	0.04	0.12	90	95
2	3.78	7.93	49	55	6.56	38.39	1.76	7.98	74	79
3	0.48	0.98	82	85	0.93	5.19	0.04	0.17	96	97
4	0.53	1.18	70	77	0.42	3.63	0.05	0.28	88	93
5	3.32	17.55	−16	19	1.10	63.50	1.30	34.00	7	42
6	2.19	4.19	58	64	4.47	22.61	0.79	2.73	83	88
7	0.78	1.20	78	83	1.72	6.98	0.08	0.25	95	96
8	1.15	1.95	75	82	2.61	15.14	0.16	0.53	94	96
9	0.37	1.32	81	86	0.54	12.22	0.03	0.35	94	97
10	0.65	1.95	79	86	2.27	49.12	0.10	0.91	95	98
11	2.42	6.14	73	76	11.45	94.46	0.78	4.35	93	95
12	2.54	5.49	64	68	6.47	39.54	0.83	3.35	86	91
13	0.19	2.48	73	81	0.06	19.93	0.02	0.83	72	96
14	2.15	3.13	70	74	9.88	26.37	0.71	1.44	93	94
15	2.00	2.62	72	77	7.45	18.88	0.48	0.84	93	96
16	0.53	3.37	57	72	0.20	18.24	0.04	1.14	79	94
17	0.89	1.72	80	86	2.69	20.96	0.11	0.47	96	98
18	2.21	2.97	65	70	5.92	14.16	0.63	1.20	89	92
19	0.79	2.54	35	61	0.43	14.12	0.19	1.98	67	87
20	0.49	2.81	54	69	0.16	11.79	0.04	0.97	76	92
Median	0.84	2.58	71	77	2.00	18.56	0.14	0.94	89	95
25-prctile	0.52	1.62	58	69	0.44	12.12	0.04	0.44	78	91
75-prctile	2.19	3.58	76	83	6.06	29.38	0.73	2.16	94	96

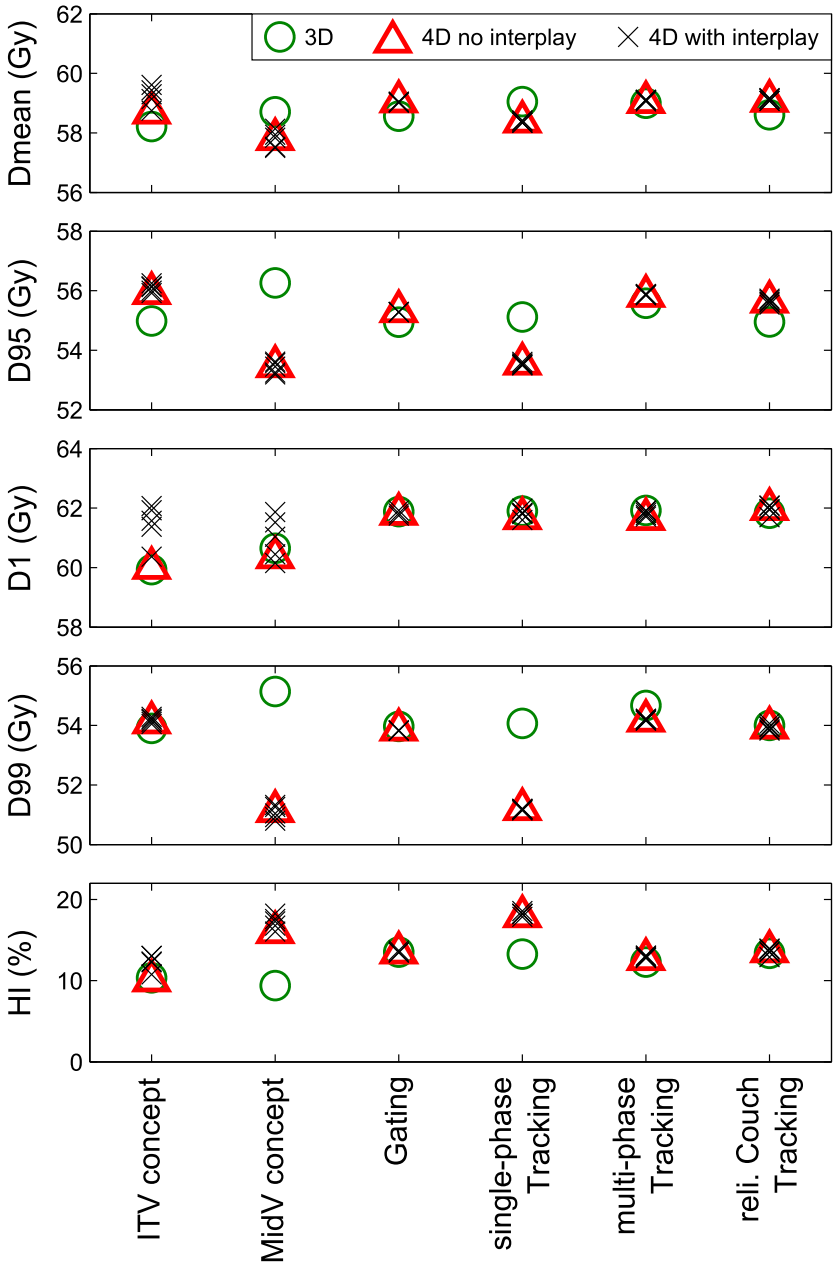
Res.: residual, ACR/ECR: amplitude/error compensation ratio, RSME: root mean square error
prctile: percentile

5 Dependence on starting phase for 4D dose calculation

In the 4D dose calculation process, the arc segments were assigend to the different respiratory phases to account for interplay effects. The assignment starts at one specific phase. The result of the 4D calculation with interplay might therefore depend on the choice of this starting phase. The dependence on this starting phase was investigated for patient 6 with a tumor motion amplitude of 13 mm. The 4D dose to the tumor was calculated for each technique with five different starting phases, except for gating, where only the three phases in the gating window were considered as starting phases. Also a 4D calculation neglecting the interplay effect was performed for each concept by simply assigning the whole plan to each phase but only with a proportion of the monitor units. The results are shown in Figure 4. For the gating and tracking treatments, no relevant dependence on the starting phase was found (spread <0.4 Gy). In these cases, the 4D dose could simply be estimated without regarding interplay effects. But for ITV and MidV concept, the mean and maximal dose (Dmean and D1) showed a spread of about 0.8 Gy and 1.7 Gy, respectively, depending on the starting phase. However, the minimal dose (D99), which gives the coverage of the target volume, was barely affected by the interplay effect itself.



Supplementary Figure 3: Target dose for gating. The Dmean, D95, D1 and HI of the TV from the original treatment plans (3D: gray boxplots) and the accumulated GTV doses (4D: white boxplots). $4D_{Mid}$: accumulation against initial countouring phase inbetween inhale and exhale. $4D_{EOI}$: accumulation against mid-gating phase at end of inhale. Arrows indicate significant differences.



Supplementary Figure 4: Dependence of 4D dose calculation on the starting phase for the arc-segment assignment. The 3D dose parameters (circles) of the target volume and the 4D dose parameters of the GTV with (crosses) and without (triangles) consideration of interplay are shown for all motion-mitigation techniques. Dmean: mean dose, D95/1/99: dose to 95/1/99% of volume, HI: homogeneity index.

References

[1] M. van Herk, P. Remeijer, C. Rasch, and J. V. Lebesque, "The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 47, no. 4, pp. 1121–1135, 2000.

[2] J. W. Wolthaus, J.-J. Sonke, M. van Herk, J. S. Belderbos, M. M. Rossi, J. V. Lebesque, and E. M. Damen, "Comparison of different strategies to use four-dimensional computed tomography in treatment planning for lung cancer patients," *International Journal of Radiation Oncology* Biology* Physics*, vol. 70, no. 4, pp. 1229–1238, 2008.

[3] S. Lang, B. Shrestha, S. Graydon, F. Cavelaars, C. Linsenmeier, J. Hrbacek, S. Klöck, G. Studer, and O. Riesterer, "Clinical application of flattening filter free beams for extracranial stereotactic radiotherapy," *Radiotherapy and Oncology*, vol. 106, no. 2, pp. 255–259, 2013.

[4] S. Lang, J. Zeimetz, G. Ochsner, M. S. Daners, O. Riesterer, and S. Klöck, "Development and evaluation of a prototype tracking system using the treatment couch," *Medical Physics*, vol. 41, no. 2, p. 021720, 2014.